Construction of lactic acid overproducing Clostridium thermocellum through enhancement of lactate dehydrogenase expression

Elsevier Editorial System(tm) for Enzyme and Microbial Technology

Manuscript Draft

Manuscript Number:

Title: Construction of lactic acid overproducing Clostridium thermocellum through enhancement of lactate dehydrogenase expression

Article Type: Research Paper

Keywords: metabolic engineering; anaerobic bacteria; ethanol; transcriptional promoter; lignocellulose

Corresponding Author: Dr. Roberto Mazzoli, Ph.D.

Corresponding Author's Institution: University of Torino

First Author: Roberto Mazzoli, Ph.D.

Order of Authors: Roberto Mazzoli, Ph.D.; Daniel G Olson; Lee R Lynd

Manuscript Region of Origin: ITALY

Abstract: Rapid expansion of global market of lactic acid (LA) has prompted research towards cheaper and more eco-friendly strategies for its production. Nowadays, LA is produced mainly through fermentation of simple sugars or starchy biomass (e.g. corn) and its price is relatively high. Lignocellulose could be an advantageous alternative feedstock for LA production owing to its high abundance and low cost. However, the most effective natural producers of LA cannot directly ferment lignocellulose. So far, metabolic engineering aimed at developing microorganisms combining efficient LA production and cellulose hydrolysis has been generally based on introducing designer cellulase systems in natural LA producers. In the present study, the approach consisted in improving LA production in the natural cellulolytic bacterium Clostridium thermocellum DSM1313. The expression of the native lactate dehydrogenase was enhanced by functional replacement of its original promoter with stronger ones resulting in a 10-fold increase in specific activity, which resulted in a 2-fold increase of LA yield. It is known that eliminating allosteric regulation can also increase lactic acid production in C. thermocellum, however we were unable to insert strong promoters upstream of the de-regulated ldh gene. A strategy combining these regulations and inactivation of parasitic pathways appears essential for developing a homolactic C. thermocellum.

Suggested Reviewers: Michael WW Adams
Biochemistry and Molecular Biology, University of Georgia
adamsm@uga.edu
Expert of engineering lactic acid production in cellulolytic microorganisms

Itzhak Mizrahi
Life Sciences, Ben-Gurion University of the Negev
imizrahi@bgu.ac.il
Expertize in cellulolytic bacteria and lactic acid production

Regiane Alves de Oliveira
Chemical engineering, Universidade Estadual de Campinas Faculdade de Engenharia Quimica
raoliveira@feq.unicamp.br
Expertize in industraial production of lactic acid from renewable feedstocks

Yixing Zhang
Grain Science and Industry, Kansas State University
yiixingzh@gmail.com
Expertize in lactic acid production from lignocellulose

Cover letter

Dear Editor,

please find enclosed the manuscript entitled "Construction of lactic acid overproducing *Clostridium thermocellum* through enhancement of lactate dehydrogenase expression". Lactic acid has several applications including the production of biodegradable plastics, namely polylactide and its co-polymers. However, a significant reduction of the cost of lactic acid production is necessary to make polylactide cost competitive with oil-derived plastics. Development of microbial strains that combine efficient metabolization of lignocellulose and biosynthesis of lactic acid by metabolic engineering can enable development of cheap single-step fermentation processes. So far, most studies have attempted to confer cellulolytic ability to native lactic acid producers. In the present study, we improved lactic acid production in one of the best cellulose degraders isolated so far, namely *Clostridium thermocellum*. By functional replacement of the original promoter of lactate dehydrogenase with stronger ones we could double lactic acid yield. As far as we know, this is the first example of improvement of lactic acid production in a native cellulolytic microorganism by targeted metabolic engineering. Apart from successful results presented here, the present investigation provided information of complex relationship between fermentative pathways in *C. thermocellum* and hints for further enhancement of lactic acid production in this strain.

We hope that the present study will be worth of interest for publication in Enzyme and Microbial Technology. Waiting for your answer we thank you very much in advance for your attention.

Yours Sincerely,

Roberto Mazzoli

Università degli Studi di Torino
DIPARTIMENTO di SCIENZE DELLA VITA E BIOLOGIA DEI SISTEMI

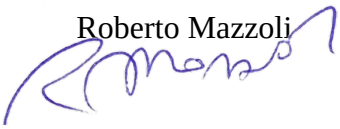
Torino, May, 28th 2020

Author agreement

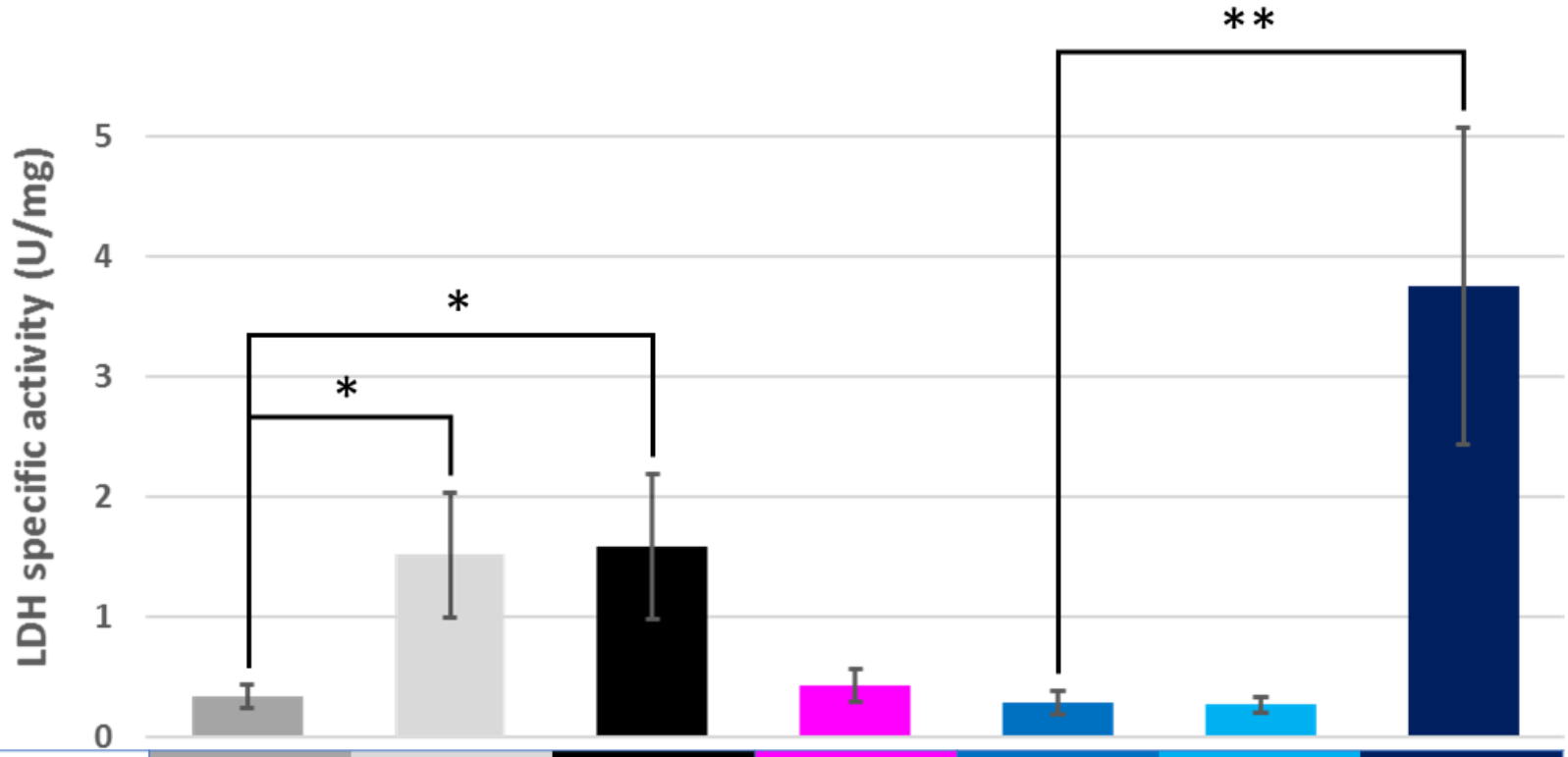
By the present letter, I do certify that all the authors have seen and approved the final version of the manuscript entitled "Construction of lactic acid overproducing *Clostridium thermocellum* through enhancement of lactate dehydrogenase expression" that has being submitted at Enzyme and Microbial Technology. I warrant that the article is the authors' original work, hasn't received prior publication and isn't under consideration for publication elsewhere.

Sincerely yours,

Roberto Mazzoli



Graphical Abstract



Strain	LL345	LL1624	LL1640	LL1111	LL1147	LL1625	LL1626
Parent strain	LL345			LL1111	LL1147		
ldh promoter	WT	P_{Thl}	P_{2638}	WT	WT	P_{Thl}	P_{2638}
Y_{LA} (mol/mol cellobiose)	0.4 (0.0)	0.8 (0.1)	0.8 (0.1)	2.5 (0.5)	0.2 (0.0)	0.0 (0.0)	0.4 (0.0)

Highlights

- Strong (P_{Thl}, P_{2638}) promoters were inserted upstream of *C. thermocellum ldh* gene
- Functional replacement of *ldh* promoter induced up to 13-fold overexpression of LDH
- Increase of LDH activity doubled lactate yield in engineered *C. thermocellum*
- LDH upregulation induced additional metabolic (ethanol, formate) rearrangements

Abstract

Rapid expansion of global market of lactic acid (LA) has prompted research towards cheaper and more eco-friendly strategies for its production. Nowadays, LA is produced mainly through fermentation of simple sugars or starchy biomass (e.g. corn) and its price is relatively high. Lignocellulose could be an advantageous alternative feedstock for LA production owing to its high abundance and low cost. However, the most effective natural producers of LA cannot directly ferment lignocellulose. So far, metabolic engineering aimed at developing microorganisms combining efficient LA production and cellulose hydrolysis has been generally based on introducing designer cellulase systems in natural LA producers. In the present study, the approach consisted in improving LA production in the natural cellulolytic bacterium *Clostridium thermocellum* DSM1313. The expression of the native lactate dehydrogenase was enhanced by functional replacement of its original promoter with stronger ones resulting in a 10-fold increase in specific activity, which resulted in a 2-fold increase of LA yield. It is known that eliminating allosteric regulation can also increase lactic acid production in *C. thermocellum*, however we were unable to insert strong promoters upstream of the de-regulated *ldh* gene. A strategy combining these regulations and inactivation of parasitic pathways appears essential for developing a homolactic *C. thermocellum*.

Construction of lactic acid overproducing *Clostridium thermocellum* through enhancement of lactate dehydrogenase expression

Mazzoli R.[a,b], Olson D.G.[b], Lynd L.R.[b]

[a]Structural and Functional Biochemistry. Laboratory of Proteomics and Metabolic Engineering of Prokaryotes. Department of Life Sciences and Systems Biology. University of Torino. Via Accademia Albertina 13. 10123 Torino. Italy.

[b]Thayer School of Engineering, Dartmouth College, 14 Engineering Drive, Hanover, NH 03755, USA.

Corresponding Author

Roberto Mazzoli

Department of Life Sciences and Systems Biology. University of Torino. Via Accademia Albertina 13. 10123 Torino. Italy.

Fax +39 011 6704508

E-mail : roberto.mazzoli@unito.it

Abstract

Rapid expansion of global market of lactic acid (LA) has prompted research towards cheaper and more eco-friendly strategies for its production. Nowadays, LA is produced mainly through fermentation of simple sugars or starchy biomass (e.g. corn) and its price is relatively high. Lignocellulose could be an advantageous alternative feedstock for LA production owing to its high abundance and low cost. However, the most effective natural producers of LA cannot directly ferment lignocellulose. So far, metabolic engineering aimed at developing microorganisms combining efficient LA production and cellulose hydrolysis has been generally based on introducing designer cellulase systems in natural LA producers. In the present study, the approach consisted in improving LA production in the natural cellulolytic bacterium *Clostridium thermocellum* DSM1313. The expression of the native lactate dehydrogenase was enhanced by functional replacement of its original promoter with stronger ones resulting in a 10-fold increase in specific activity, which resulted in a 2-fold increase of LA yield. It is known that eliminating allosteric regulation can also increase lactic acid production in *C. thermocellum*, however we were unable to insert strong promoters upstream of the de-regulated *ldh* gene. A strategy combining these regulations and inactivation of parasitic pathways appears essential for developing a homolactic *C. thermocellum*.

Key words: metabolic engineering, anaerobic bacteria, ethanol, transcriptional promoter, lignocellulose

Introduction

Lactic acid (LA) is among the most requested chemicals worldwide because of its broad range of industrial applications [1]. The latter include the use as food preservative and flavor enhancer, emulsifier and moisturizer in the cosmetic industry, precursor of pharmaceuticals and biodegradable solvents, and building block for the synthesis of plastic polymers, i.e., polylactide (PLA) and its co-polymers [2]. PLAs are biodegradable and biocompatible plastics whose application encompasses several sectors including biomedicine (e.g. surgical thread, orthopedic implants, drug delivery), packaging of food and goods, manufacturing of agriculture mulch films and disposable tableware [2,3]. The potential of PLA to replace fossil-fuel-derived polymers as a general-purpose plastic has been among the main forces driving the current global market expansion of LA. The growth of global LA demand is currently estimated at 16.2% per year [2].

LA can be produced by either chemical synthesis from acetaldehyde or by microbial fermentation. However, most (about 90%) LA production plants worldwide are based on fermentation of starchy biomass (mainly corn) [2,4]. Biotechnological production of LA has many advantages over chemical synthesis such as lower energy consumption and environmental concerns and higher purity [2]. In particular, fermentative production can lead to optically pure L- or D-LA, while a racemic mixture of the two LA enantiomers is obtained by chemical synthesis [5]. It is worth remembering that a precise mixture of D- and L-LA is required for production of PLA with desired physical-chemical characteristics [5]. Yet, for economic reasons, industrial production of PLA is considered a relatively immature technology [4]. PLA is still too expensive to compete with fossil-derived plastics and this is mainly due to the cost of LA. The current cost of LA is relatively high ($1.30-4.0/kg) and suffers from significant variations of the price of starch or sugar feedstocks used for the fermentation process [6]. In fact, the cost of the feedstock is among the most relevant parameters determining the fermentation cost. This concern has stimulated research towards utilization of alternative feedstocks such as milk whey, food waste, glycerol, or microalgae [1,2,4].

In particular, significant attention has been dedicated to lignocellulosic biomass owing to its high abundance and low cost. Lignocellulose includes most waste biomass such as agricultural/land by-products (cereal straw, sugar cane bagasse, forest residues), municipal solid wastes and industrial wastes (e.g. paper mill sludge) [7]. However, the most efficient natural producers of LA, i.e. lactic acid bacteria (LAB), bacteria belonging to the *Bacillus* genus and fungi belonging to the *Rhizopus* genus, cannot ferment lignocellulosic material without prior biomass saccharification [8,9]. Biomass pre-treatment and, in particular, exogenous cellulase supplementation are highly expensive and significantly increase the cost of the entire process thus making it not-viable from an economic standpoint [10,11]. Development of consolidated bioprocessing (CBP), that is one-pot fermentation, of lignocellulose is therefore highly desirable, as a mean to significantly lower the cost of lignocellulose fermentation to LA and make PLA cost-competitive with oil-derived plastics. Recently, an example of CBP based on an artificial consortium consisting of a cellulolytic fungus (i.e. *Trichoderma reesei*) and a LAB (i.e. *Lactobacillus pentosus*) has been reported (Shahab et al., 2018). However, industrial exploitation of this approach requires improved robustness, stability and reproducibility of co-cultures [12]. Metabolic engineering has been used to develop recombinant strains that combine high LA production and efficient biomass fermentation. Most studies have been aimed at introducing cellulolytic characteristics (e.g. by expression of heterologous cellulases) in natural LA producers, such as LAB [8]. In the present study, we used a different approach, i.e. we attempted to improve LA production in a native cellulolytic microorganism, namely *Clostridium thermocellum*. Metabolic engineering strategies addressed to native cellulolytic microorganisms have generally been focused on increasing biofuel, namely ethanol and butanol, production [13,14]. However, these studies have also indicated suitable metabolic targets for improving LA production [9].

Clostridium thermocellum DSM1313 is a thermophilic cellulolytic bacterium among the best cellulose degraders and the most promising candidates for application in CBP of plant biomass to

biofuels and other high-value chemicals. A significant amount of information on the central metabolism and, more in general, on the biology of *C. thermocellum* is currently available [15–17]. Furthermore, reliable methods for transformation, inducible gene expression and markerless gene deletion have been developed for this strain [18,19].Wild type *C. thermocellum* produces a mixture of organic acids (including acetic acid, formic acid and LA), ethanol, H_2 and CO_2. The main carbon catabolites are ethanol and acetate, while LA yield is very low (i.e. 0.01 mol/mol hexose equivalent) [20]. Improvement of the production of a chemical in a microorganism can be obtained through increased expression of enzyme(s) directly involved in its biosynthesis, and/or disruption of competing pathways [21]. Recently, dramatic increase in LA yield in *C. thermocellum* was obtained by deleting the autologous *adhE* gene, that encodes the main bifunctional alcohol/aldehyde dehydrogenase [20]. This modification almost abolished ethanol production in the engineered strain (i.e. strain LL1111) and significantly re-directed *C. thermocellum* carbon flux towards production of LA. Actually, LA is the main end-catabolite of LL1111 which is the *C. thermocellum* strain with the highest LA yield (40% of the maximum theoretical yield) obtained so far [20]. It is worth noting that LL1111 also features a spontaneously occurred mutation of its *ldh* gene resulting in a LDH whose activity is independent from allosteric activation by fructose 1,6 bisphosphate (F1,6BP) [20]. The strategy used in the present study aimed at enhancing the expression of native LDH by functionally replacing the original transcriptional promoter of the unique *ldh* gene (Clo1313_1160) [22] with stronger ones. Such modification has been performed on the "wild type" *C. thermocellum* (LL345) and on strains LL1147 [23] and LL1111 [20] that were recently engineered. Strain LL1147 is repressed in H_2 production because of functional inactivation of all four hydrogenases. Most attempts were successful and led to significant improvement (4.5-13 fold) of LDH activity and LA yield in the engineered strains.

Materials and Methods

Bacterial strains and culture conditions

All reagents used in this study were of molecular grade, and obtained either from Sigma Aldrich or Fisher Scientific, unless otherwise stated. *C. thermocellum* DSM1313 was obtained from the DSMZ culture collection. It was grown in either chemically defined MTC-5 medium at initial pH of 7.4 [24] or in rich CTFUD medium at initial pH of 7.0 [18] supplemented with 5 g/l cellobiose as the main carbon source. Cultures were incubated at 55°C under anaerobic conditions either in conical tubes in anaerobic chambers (Coy Laboratory Products, Grass Lakes, MI, USA) or in 125 ml (containing 50 ml of medium) butyl stoppered vials.

For measurement of growth parameters, strains were grown in a 96-well plate on in 200 µl of MTC-5 medium and absorbance at 600 nm was determined every 3 min for 72 h in a Powerwave XS plate reader as previously described [25]. For measurement of fermentation products, strains were grown in CTFUD medium. Samples were harvested immediately after inoculation of the medium, and after 72 h of growth. Data for fermentation products and growth rate are averages from biological triplicate experiments.

Analytical techniques

Cellobiose, glucose, acetate, citrate, formate, ethanol, lactate, malate, pyruvate and succinate were measured by HPLC using an Aminex HPX-87H column (BioRad, CA, USA) equipped with both refractive index and UV detector as previously reported [26].

Gene modification

Gene modification was performed in the genetically tractable Δ*hpt* strain of *C. thermocellum* DSM1313 (referred as LL345) [22], and two strains derived from LL345 with additional deletion

of: i) *hydG* and *ech* genes (i.e. strain LL1147), encoding HydG, involved in maturation of the three [Fe-Fe] hydrogenases, and the [Ni-Fe] hydrogenase, respectively [23]; ii) *adhE* (i.e. strain LL1111) encoding the main aldehyde/alcohol dehydrogenase [20].

Gene modification consisted in insertion of the thiolase (*thlA*) promoter (P_{Thl}) of *Clostridium acetobutylicum* [27] or the promoter of Clo1313_2638 (P_{2638}) of *C. thermocellum* DSM1313 [28] upstream (position 1,381,634 of *C. thermocellum* chromosome) of *ldh* gene (Clo1313_1160) of *C. thermocellum*. The inserted P_{Thl} sequence corresponded to nucleotide 2450 to 2612 of the shuttle vector pSOS95 (GenBank accession number: AY187686.1), while P_{2638} included nucleotide 3,106,958 to 3,106,750 of *C. thermocellum* chromosome. Two gBlocks (Integrated DNA Technologies), one containing the 500 bp fragment upstream (5' flank) of *ldh* gene and the other consisting of 500 bp upstream (5' flank) of *ldh* gene, the promoter of choice and the first 500 bp of *ldh* gene (3' flank), were designed (Supplementary Table S1). The gBlocks were cloned into the pDGO145 plasmid (GenBank accession number: KY852359) in the EcoRV and PvuII sites by Gibson assembly so as to obtain the plasmids for promoter integration into *C. thermocellum* chromosome [18]. The plasmids were transformed in T7 express chemiocompetent *E. coli* cells to ensure proper methylation [29]. Methylated plasmids were transformed into *C. thermocellum* by electroporation as described previously [18]. Integration of P_{Thl} and P_{2638} in the suited genome locus was obtained for *C. thermocellum* strain LL345 (thus resulting in strains LL1624 and LL1640, respectively) and LL1147 (thus resulting in strains LL1625 and LL1626, respectively) but not in strain LL1111. Genetic modifications were confirmed by whole-genome resequencing by the Department of Energy Joint Genome Institute. Raw genome resequencing data are available from the NCBI Sequence Read Archive (SRA; http://www.ncbi.nlm.nih.gov/Traces/sra). Data were analyzed with the CLC Genomic Workbench version 11.0.1, (Qiagen Inc., Hilden, Germany), as previously described [26]. Strains used in this study are listed in Table 1.

Enzyme assays

Cells were grown in CTFUD medium to an optical density at 600 nm (OD_{600nm}) of 0.6 and harvested by centrifugation (7,000 x g, 10 min, 4°C). Harvested cells were washed twice in anaerobic conditions with cold 100 mM Tris-HCl pH 7 (12,000 x g, 5 min) and stored at -80°C. Protein extracts were prepared by incubating cells with lysozyme as previously described [20]. Protein content was measured by Bio-Rad protein dye reagent with bovine serum albumin (Thermo Scientific) as the standard. Lactate dehydrogenase (LDH) activity assays were performed as previously described [20]. Briefly, assays were performed at 55°C in a Coy anaerobic chamber with an 85% N_2, 10% CO_2, and 5% H_2 atmosphere maintained under anoxic conditions using a palladium catalyst. NADH consumption was monitored by measuring absorbance at 340 nm using an extinction coefficient of 6,220 M^{-1} cm^{-1}. The reaction conditions for lactate dehydrogenase (LDH) activity were 200 mM Tris-HCl (pH 7.3), 0.22 mM NADH, 10 mM sodium pyruvate, and 1 mM fructose 1,6-bisphosphate (F1,6BP). The reaction was started by addition of sodium pyruvate.

Results

Functional replacement of ldh promoter in C. thermocellum

Plasmids for inserting the promoter (P_{Thl}) of thiolase gene (*thlA*) from *C. acetobutylicum* or of the promoter (P_{2638}) of Clo1313_2638 of *C. thermocellum* DSM1313 upstream of the *ldh* gene (Clo1313_1160) of *C. thermocellum* DSM1313 were constructed and transformed in *C. thermocellum* strains LL345, LL1111 and LL1147. Both promoters are known for being strong and constitutive [27,28]. LL345 was obtained by disruption of the *hpt* gene in *C. thermocellum* DSM1313, but can considered as the wild-type strain from a metabolic standpoint, since the *hpt* gene was deleted to allow for 8AZH counter-selection, and it does not affect the fermentation phenotype [22]. LL1111 and LL1147 have been obtained from LL345 respectively by: i) deletion of

148 the gene encoding the main bifunctional alcohol/aldehyde dehydrogenase AdhE (Lo et al., 2015); ii)
149 disruption of *hydG*, encoding maturase of the three [FeFe] hydrogenases, and *ech*, encoding the
150 [NiFe] hydrogenase, resulting in a *C. thermocellum* strain that lacks hydrogenase activity (Biswas et
151 al., 2015). In the present study, functional replacement of *ldh* promoter with P_{Thl} has been obtained
152 in the *C. thermocellum* strains LL345 (wt) and LL1147 (hydrogenase knockout), resulting in strains
153 LL1624 and LL1625, respectively. Insertion of P_{2638} in strains LL345 and LL1147 resulted in
154 strains LL1640 and LL1626, respectively. Attempts to introduce P_{Thl} or P_{2638} upstream of the *ldh*
155 gene in strain LL1111 (*adhE* deletion) were unsuccessful. Genome re-sequencing of all the
156 engineered strains confirmed successful modification of the *ldh* locus in strains LL1624, LL1625,
157 LL1626 and LL1640.

158

159 ***LDH activity of engineered strains***

160 LDH activity in parent and engineered strains was measured (Figure 1). Furthermore, LDH assays
161 were performed on strain LL1111 (*adhE* deletion) since it is the *C. thermocellum* strain with the
162 highest LA production yield (40% of the maximum theoretical yield) obtained so far (Lo et al.,
163 2015). In three out of the four strains engineered in this study (i.e. LL1624, LL1626 and LL1640),
164 specific LDH activity was significantly improved with respect to their parent strains by a factor
165 ranging from 4.5 to 13-fold (Figure 1). This confirms that the promoters chosen were stronger than
166 the original *ldh* promoter. With respect to their parent strain, namely LL345, strains LL1624 and
167 LL1640 showed about 4.5-fold increased LDH specific activity, which suggests that P_{Thl} and P_{2638}
168 have similar strength. However, these promoters did not have the same effect in the LL1147
169 (hydrogenase deletion) background. The highest LDH specific activity was measured in LL1626,
170 that is LL147 with P_{2638}, with an increase of more than 13-fold with respect to the parent strain
171 LL1147. Unexpectedly, no enhancement of LDH activity was detected in LL1625 carrying the P_{Thl}
172 in LL1147 genetic background (Figure 1).

Growth and fermentation profiles of engineered strains

Overexpression of LDH had little effect on growth of recombinant *C. thermocellum* strains obtained in the present study (Table 2). Most frequently, a slight (ranging from 26 to 29%) reduction in the specific growth rate was observed, i.e. in strains LL1624 and LL1626, with respect to parent strains. An even weaker reduction (8-9 %) of final biomass was shown by strains LL1626 and LL1640 (with respect to parent strain LL1147 (hydrogenase deletion) and LL345 (wt)). Although functional replacement of the *ldh* promoter did not increase LDH activity, strain LL1625 showed a significantly lower growth rate (17%) and slightly increased final biomass with respect to LL1147.

Determination of substrate and fermentation product concentrations was performed immediately after inoculation, and then 72 hours later, after growth had stopped (Table 2). Cellobiose consumption was largely unchanged by the introduction of different promoters driving the *ldh* gene. In several cases, however, overexpression of *ldh* resulted in less glucose production (strains LL1624, LL1625, and LL1626).

Apart from sugar consumption, also the profile of fermentation end-products of the engineered strains was affected. In strains where higher LDH activity was measured, also LA yield was improved although to a lower extent. In the wild type strain background (LL345), overexpression of LDH approximately doubled the LA yield (strains LL1624 and LL1640), consistent with enzyme assay data. In the hydrogenase deletion background (LL1147), overexpression of LDH had a variable effect on LA yield. The P_{Thl} promoter decreased LA yield (LL1625), while the P_{2638} promoter increased (about 2-fold) LA yield (LL1626). Maximum LA yield obtained in this study (i.e. 0.8 mol/mol cellobiose, strain LL1640) is still dramatically lower than that of LL1111, i.e. the *C. thermocellum* strain with the highest LA yield obtained so far (Lo et al., 2015). Apart from disruption of *adhE* gene, in strain LL1111 a spontaneous mutation of the *ldh* gene appeared,

resulting in a mutant LDH having high catalytic activity even without its allosteric activator, namely F1,6BP [20].

Overexpression of LDH induces other re-arrangements of the metabolic network of *C. thermocellum*, and the overall fermentation profile of the engineered strains is significantly different from that of the parent strains. Most frequently, improved LA yield is accompanied by increased ethanol yield (strains LL1624 and LL1626) and/or reduction of formate yield (strains LL1640 and LL1626) (Table 2). Interestingly, catabolite profiles of LL1624 and LL1640 differ from each other, although they show similar LDH activity and LA yield and both derive from LL345. LL1624 shows a significant reduction of acetate yield and an increase of ethanol and malate production. In LL1640, only formate yield resulted as significantly decreased with respect to LL345. Metabolite profiles of strains derived from LL1147 were more similar between each other. Both LL1625 and LL1626 showed decreased production of formate and malate and remarkable (more than 2-fold) increase in ethanol yield. It is worth noting that LL1626 has the highest LDH specific activity measured in this study and its LA yield is about 2-fold higher than that of LL1647, but LL1625 has LDH levels similar to the parent strains.

Genome sequencing of engineered strains

After strain construction, correct insertion of the promoter upstream of *ldh*, as well as the absence of point mutations in or upstream of the *ldh* gene was verified by whole genome resequencing. However, there were several unexpected results from the enzyme assay and fermentation data, so we looked to see if these results could be explained by any of the secondary mutations we observed. The LDH enzyme activity in strain LL1625 was no different from that of its parent strain, LL1147, despite the presence of the P_{Thl} promoter. We identified seven mutations that were present LL1625, but not in its parent strain (LL1147) (Figure 2). Of these mutations, the one starting at position 279492 (coordinates based on Genbank sequence NC_017304.1) should be ignored because it is

present in both strains LL1625 and LL1626. Of the remaining six mutations, we do not have any direct evidence to favor one over another. There are two which deserve additional scrutiny. There is a point mutation in the Clo1313_1324 gene. This gene is thought to encode *rpoD*, which is the sigma factor associated with basal expression (i.e. "housekeeping" genes), and could affect transcription from the P_{Thl} promoter. There is also a mutation in the Clo1313_1122 start codon that changed it from an ATG to a GTG. This type of mutation is typically associated with reduced transcription. This is interesting because in LL1147, the parent strain of LL1625, this position was mutated from GTG to ATG, so in strain LL1625, the mutation reverted back to the wild type.

Another unexplained phenotype is the increased ethanol production and decreased acetate production of strain LL1624 relative to either its parent (LL345) or sibling (LL640) strains. Although there are 17 potential mutations that are unique to strain LL1624, none of them are obviously associated with ethanol production, acetate production, or redox balance.

Discussion

Production of LA through a 2nd generation biorefinery approach has attracted significant interest because of the economic advantages that using lignocellulose as the fermentation feedstock could bring on LA price [1,2]. The present study aimed at improving LA production in the native cellulolytic bacterium *C. thermocellum* by enhancing the expression of its LDH. The final purpose of this investigation is developing a strain able to catalyze direct fermentation of lignocellulose to LA. Development of such CBP could significantly reduce the current cost of LA. Most previous metabolic engineering strategies aimed at direct fermentation of lignocellulose to LA have attempted to introduce (hemi)cellulolytic characteristics in natural producers of LA [30,31]. The approach used in the present study focused on improving the expression of the native LDH of the cellulolytic bacterium *C. thermocellum* DSM1313 by functionally replacing its native promoter

with stronger ones. To this aim, two strong and constitutive promoters, i.e. the one of thiolase (P_{Thl}) from *C. acetobutylicum* [27] and that of Clo1313_2638 (P_{2638}) from *C. thermocellum* DSM1313 [28], were used. In three out of the four strains engineered in this study (i.e. LL1624, LL1626 and LL1640), functional replacement of the *ldh* promoter led to significant improvement of specific LDH activity with respect to parent strains by a factor ranging from 4.5 to 13-fold (Figure 1). These strains show the highest specific LDH activities reported in *C. thermocellum*, so far, that is about 4-9 fold higher than that measured in strain LL1111 (i.e. the *C. thermocellum* strain with the highest LA yield reported so far). It is worth remembering that, in many microorganisms, LDH expression is under the control of the global redox-responsive transcription factor Rex [32]. Generally, Rex acts as a gene transcription repressor in response to low intracellular [NAD(P)H]/[NAD(P)+] ratio. Although this has not been confirmed in *C. thermocellum* yet, *ldh* promoter engineering obtained in this study may have altered this regulation. Recently, a similar approach was used to improve the expression of the *ldh* gene from *Caldicellulosyruptor bescii* [33]. In this case the original promoter was replaced with the xylose-inducible promoter P_{xi} but improvement of the specific LDH activity with respect to the wild type strain (about 3 fold) was lower than that obtained in the present study.

However, the effect of *ldh* promoter replacement on LA production was milder since no more than 2-fold increase of LA yield with respect to parent strains was observed. Maximum LA yield obtained in this study (i.e. 0.8 mol/mol cellobiose, strain LL1640) is still dramatically lower than that of strain LL1111 (Lo et al., 2015). Apart from disruption of *adhE* gene, in strain LL1111 a spontaneous mutation of the *ldh* gene appeared, resulting in a mutant LDH having high catalytic activity even without its allosteric activator, namely F1,6BP [20]. LA yield of strain LL1626 (derived from the hydrogenase-deficient strain LL1147) was even lower despite this strain shows the highest LDH specific activity (Figure 1). This observation is most probably related to the different metabolic background of this strain. Strain LL1147 also shows dramatic reduction of LA accumulation [23]. The exact cause of this metabolic phenotype was not determined but it was

speculated that disruption of hydrogenases could have altered intracellular levels of possible allosteric regulators of LDH [23]. Apart from the abovementioned F1,6BP, the LDH enzymes may also be activated by ATP and inhibited by pyrophosphate, e.g. in *Caldicellulosiruptor saccharolyticus* [34]. Nicotinamide cofactors are other typical regulators of LDH activity such as in *Caldicellulosiruptor saccharolyticus*, where NAD^+ is a competitive inhibitor [34], or in *Thermoanaerobacter ethanolicus* where, curiously, LDH is inhibited by NADPH [35]. It is likely that hydrogenase-deleted *C. thermocellum* features accumulation of reduced ferredoxin via PFOR which could cause accumulation of other reduced electron carriers such as NADPH possibly leading to inhibition of LDH [23].

The moderate increase of LA yield obtained by the present study was not completely unexpected. Overexpression of an enzyme, although it can significantly divert the metabolism towards the product of interest, as in this case, generally, is not sufficient for driving all the carbon flux towards the pathway of interest. More in detail, previous studies have demonstrated that lactate production in C. *thermocellum* is affected at multiple levels, i.e. by allosteric regulation and deletion of competing pathways [20]. The present study has identified a third factor: transcriptional regulation, since introducing stronger promoters upstream of the *ldh* gene increases both LDH enzyme activity and LA production. All these factors contribute to sophisticated regulation of LA production in *C. thermocellum*. A previous study has shown that the deletion of *adhE* likely have the largest impact on lactate production [20]. Allosteric control and transcriptional control appear to have similar strength based on similar LA yields of strains LL1624 or LL1640 compared to strain LL1160 carrying only the S161R mutation in LDH [20].

Actually, current information on *C. thermocellum* metabolism still have important gaps as regards LA production. It is known that K_m for pyruvate *C. thermocellum* LDH is highly affected by F1,6BP concentration [36], but regulation by further allosteric effectors (ATP, pyrophosphate, nicotinamide cofactors) has been hypothesized [23]. The affinity of the other *C. thermocellum*

296 enzymes that directly compete for the same substrates, that is pyruvate and/or NADH, is currently
297 not known. Alternative reactions for dissimilation of pyruvate in *C. thermocellum* are catalyzed by
298 pyruvate ferredoxin oxidoreductase (PFOR) (E.C. 1.2.7.1) and by pyruvate formate lyase (PFL)
299 (E.C. 2.3.1.54). *C. thermocellum* genome harbors five genes or gene clusters annotated as encoding
300 PFOR [37]. Some evidence indicates that *pfor1* (Clo1313_0020-0023) and *pfor4* (Clo1313_1353-
301 1356) are the primary PFOR of *C. thermocellum* [37–39]. Actually, deletion of each of these gene
302 clusters causes about 80% reduction of PFOR activity in this strain [37]. However, the K_m for
303 pyruvate of these enzymes has not been determined. *C. thermocellum* PFL is encoded by *pflB* gene
304 (Clo1313_1717) but its K_m for pyruvate has not been measured.

305 Overexpression of LDH induces other, sometimes unexpected, re-arrangements of the metabolic
306 network and the overall fermentation profile of *C. thermocellum*. In two (strains LL1624 and
307 LL1626) out of three strains with increased LA yield, also ethanol yield was enhanced and/or
308 formate yield was diminished (strains LL1640 and LL1626) (Table 2). Additionally, strain LL1624
309 shows a significant reduction of acetate yield. It is worth remembering that acetate, formate and
310 ethanol biosynthesis compete with LA production. Inhibition of formate and acetate in LA
311 overproducing strains was therefore expected. However, improvement of ethanol yield in strains
312 LL1624 and LL1626 was surprising since ethanol production competes with LA synthesis for both
313 carbon intermediates and electrons. However, a number of observations made by previous and
314 present study may explain the beneficial effect of improvement of LDH activity on fermentation
315 efficiency and, in particular, ethanol production in *C. thermocellum*. Most probably enhancement of
316 LDH activity affects the redox balance of the cells and, in particular, lowers the $NADH/NAD^+$ ratio,
317 which should be particularly high in strain LL1147, since it lacks hydrogenases. High
318 $NADH/NAD^+$ ratios have been shown to inhibit the GAPDH reaction [40,41] and reduce the
319 glycolytic flux in *C. thermocellum* [42]. Improvement of LDH activity should benefit glycolytic
320 flux because it consumes both NADH and pyruvate (as demonstrated by the fact that strain LL1626

does not show any pyruvate accumulation while its parent strain LL1147 does, Table 2). Regarding improvement of ethanol yield in strains LL1624 and LL1626, it is worth reminding that, recently, the activity of one of the main PFOR of *C. thermocellum* (i.e. PFOR1) was shown to be inhibited by NADH accumulation [43]. Furthermore, thermodynamic analysis indicated that while pyruvate dissimilation through LDH or PFL reactions is favorable even in presence of high concentrations of lactate or formate, respectively, the PFOR reaction becomes less favorable in highly reduced conditions [44]. So, it can be speculated that improved LDH activity helps removing possible inhibition of PFOR deriving from excess of reduced co-factors, and eventually this may lead to improved conversion of acetyl-CoA to ethanol.

Future perspectives to improve pyruvate flux towards LA in *C. thermocellum* should aim at improving LDH affinity for pyruvate by either: i) re-introducing the S161R mutation of *C. thermocellum* LDH as found in strain LL1111 (*adhE* deletion) or; replacing *C. thermocellum* original LDH with heterologous thermophilic LDH with higher affinity for pyruvate. As regards the first strategy, i.e. introducing a strong promoter upstream of the mutant *ldh* from LL1111, attempts to create this strain failed so far. Future attempts to create such a strain would benefit from improved genetic tools, such as tightly repressed inducible promoters. A possible candidate for the second strategy is LDH from *Thermus caldophilus* which shows a K_m for pyruvate which is 10-fold lower than that of *C. thermocellum* LDH [45].

Funding

R.M. was supported by an Italy-U.S. Fulbright Research Scholarship. Funding for D.G.O, L.R.L, and materials was provided by the Center for Bioenergy Innovation, a U.S. Department of Energy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. Resequencing was performed by the Department of Energy Joint Genome

Institute, a DOE Office of Science User Facility, and is supported by the Office of Science of the U.S. Department of Energy under contract number DE-AC02–05CH11231.

Conflict of interest

L.R.L. is a founder of the Enchi Corporation, which has a financial interest in *Clostridium thermocellum*. R.M and D.G.O declare no conflict of interest.

Author contributions

R.M., D.G.O. and L.R.L. conceived the experimental design. R.M. performed the experiments. D.G.O. analyzed re-sequencing results and contributed to metabolite analysis. All the Authors contributed to discussing the results and writing the manuscript.

Compliance with Ethical Standards

This article does not contain any studies with human participants or animals performed by any of the authors.

References

[1] R. Mazzoli, F. Bosco, I. Mizrahi, E.A. Bayer, E. Pessione, Towards lactic acid bacteria-based biorefineries, Biotechnol. Adv. 32 (2014) 1216–1236. https://doi.org/10.1016/j.biotechadv.2014.07.005.

[2] R. Alves de Oliveira, A. Komesu, C.E. Vaz Rossell, R. Maciel Filho, Challenges and

opportunities in lactic acid bioprocess design—From economic to production aspects, Biochem. Eng. J. 133 (2018) 219–239. https://doi.org/10.1016/j.bej.2018.03.003.

[3] M.A. Abdel-Rahman, Y. Tashiro, K. Sonomoto, Lactic acid production from lignocellulose-derived sugars using lactic acid bacteria: Overview and limits, J. Biotechnol. 156 (2011) 286–301. https://doi.org/10.1016/j.jbiotec.2011.06.017.

[4] M.A. Abdel-Rahman, Y. Tashiro, K. Sonomoto, Recent advances in lactic acid production by microbial fermentation processes, Biotechnol. Adv. 31 (2013) 877–902. https://doi.org/10.1016/j.biotechadv.2013.04.002.

[5] M.A. Abdel-Rahman, K. Sonomoto, Opportunities to overcome the current limitations and challenges for efficient microbial production of optically pure lactic acid, J. Biotechnol. 236 (2016) 176–192. https://doi.org/10.1016/j.jbiotec.2016.08.008.

[6] M.J. Biddy, C.J. Scarlata, C.M. Kinchin, Chemicals from biomass: A market assessment of bioproducts with near-term potential, NREL Rep. (2016) 10.2172/1244312. https://doi.org/10.2172/1244312.

[7] R.E.H. Sims, W. Mabee, J.N. Saddler, M. Taylor, An overview of second generation biofuel technologies, Bioresour. Technol. 101 (2010) 1570–1580. https://doi.org/10.1016/j.biortech.2009.11.046.

[8] L. Tarraran, R. Mazzoli, Alternative strategies for lignocellulose fermentation through lactic acid bacteria: The state of the art and perspectives, FEMS Microbiol. Lett. 365 (2018) doi: 10.1093/femsle/fny126. https://doi.org/10.1093/femsle/fny126.

[9] R. Mazzoli, Metabolic engineering strategies for consolidated production of lactic acid from lignocellulosic biomass, Biotechnol. Appl. Biochem. (2019) doi: 10.1002/bab.1869.

[10] K. Okano, T. Tanaka, C. Ogino, H. Fukuda, A. Kondo, Biotechnological production of

389 enantiomeric pure lactic acid from renewable resources: Recent achievements, perspectives,
390 and limits, Appl. Microbiol. Biotechnol. 85 (2010) 413–423. https://doi.org/10.1007/s00253-
391 009-2280-5.

392 [11] G. Liu, J. Sun, J. Zhang, Y. Tu, J. Bao, High titer l-lactic acid production from corn stover
393 with minimum wastewater generation and techno-economic evaluation based on Aspen plus
394 modeling, Bioresour. Technol. 198 (2015) 803–810.
395 https://doi.org/10.1016/j.biortech.2015.09.098.

396 [12] Y. Jiang, R. Wu, J. Zhou, A. He, J. Xu, F. Xin, W. Zhang, J. Ma, M. Jiang, W. Dong, Recent
397 advances of biofuels and biochemicals production from sustainable resources using co-
398 cultivation systems, Biotechnol. Biofuels. 12 (2019) 155. https://doi.org/10.1186/s13068-
399 019-1495-7.

400 [13] J. Lo, D.G. Olson, S.J.L. Murphy, L. Tian, S. Hon, A. Lanahan, A.M. Guss, L.R. Lynd,
401 Engineering electron metabolism to increase ethanol production in Clostridium
402 thermocellum, Metab. Eng. 39 (2017) 71–79. https://doi.org/10.1016/j.ymben.2016.10.018.

403 [14] T. Bao, J. Zhao, J. Li, X. Liu, S.T. Yang, n-Butanol and ethanol production from cellulose by
404 Clostridium cellulovorans overexpressing heterologous aldehyde/alcohol dehydrogenases,
405 Bioresour. Technol. 285 (2019) 121316. https://doi.org/10.1016/j.biortech.2019.121316.

406 [15] Y. Deng, D.G. Olson, J. Zhou, C.D. Herring, A. Joe Shaw, L.R. Lynd, Redirecting carbon
407 flux through exogenous pyruvate kinase to achieve high ethanol yields in Clostridium
408 thermocellum, Metab. Eng. 15 (2013) 151–158.
409 https://doi.org/10.1016/j.ymben.2012.11.006.

410 [16] J. Zhou, D.G. Olson, D.A. Argyros, Y. Deng, W.M. van Gulik, J.P. van Dijken, L.R. Lynd,
411 Atypical glycolysis in Clostridium thermocellum, Appl. Environ. Microbiol. (2013).
412 https://doi.org/10.1128/AEM.04037-12.

[17] D.G. Olson, M. Hörl, T. Fuhrer, J. Cui, J. Zhou, M.I. Maloney, D. Amador-Noguez, L. Tian, U. Sauer, L.R. Lynd, Glycolysis without pyruvate kinase in Clostridium thermocellum, Metab. Eng. (2017). https://doi.org/10.1016/j.ymben.2016.11.011.

[18] D.G. Olson, L.R. Lynd, Transformation of clostridium thermocellum by electroporation, in: Methods Enzymol., 2012: pp. 317–330. https://doi.org/10.1016/B978-0-12-415931-0.00017-3.

[19] E.B. Mearls, D.G. Olson, C.D. Herring, L.R. Lynd, Development of a regulatable plasmid-based gene expression system for Clostridium thermocellum, Appl. Microbiol. Biotechnol. (2015). https://doi.org/10.1007/s00253-015-6610-5.

[20] J. Lo, T. Zheng, S. Hon, D.G. Olson, L.R. Lynd, The bifunctional alcohol and aldehyde dehydrogenase gene, adhE, is necessary for ethanol production in Clostridium thermocellum and Thermoanaerobacterium saccharolyticum, J. Bacteriol. 197 (2015) 1386–1393. https://doi.org/10.1128/JB.02450-14.

[21] R. Mazzoli, Development of microorganisms for cellulose-biofuel consolidated bioprocessings: Metabolic engineers' tricks, Comput. Struct. Biotechnol. J. 3 (2012) e201210007. https://doi.org/10.5936/csbj.201210007.

[22] D.A. Argyros, S.A. Tripathi, T.F. Barrett, S.R. Rogers, L.F. Feinberg, D.G. Olson, J.M. Foden, B.B. Miller, L.R. Lynd, D.A. Hogsett, N.C. Caiazza, High ethanol Titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes, Appl. Environ. Microbiol. 77 (2011) 8288–8294. https://doi.org/10.1128/AEM.00646-11.

[23] R. Biswas, T. Zheng, D.G. Olson, L.R. Lynd, A.M. Guss, Elimination of hydrogenase active site assembly blocks H2 production and increases ethanol yield in Clostridium thermocellum, Biotechnol. Biofuels. 8 (2015) 20. https://doi.org/10.1186/s13068-015-0204-4.

[24] S. Hon, D.G. Olson, E.K. Holwerda, A.A. Lanahan, S.J.L. Murphy, M.I. Maloney, T. Zheng,

B. Papanek, A.M. Guss, L.R. Lynd, The ethanol pathway from Thermoanaerobacterium saccharolyticum improves ethanol production in Clostridium thermocellum, Metab. Eng. 42 (2017) 175–184. https://doi.org/10.1016/j.ymben.2017.06.011.

[25] S.A. Tripathi, D.G. Olson, D.A. Argyros, B.B. Miller, T.F. Barrett, D.M. Murphy, J.D. McCool, A.K. Warner, V.B. Rajgarhia, L.R. Lynd, D.A. Hogsett, N.C. Caiazza, Development of pyrF-Based genetic system for targeted gene deletion in clostridium thermocellum and creation of a pta mutant, Appl. Environ. Microbiol. 76 (2010) 6591–6599. https://doi.org/10.1128/AEM.01484-10.

[26] J. Zhou, D.G. Olson, A.A. Lanahan, L. Tian, S.J.L. Murphy, J. Lo, L.R. Lynd, Physiological roles of pyruvate ferredoxin oxidoreductase and pyruvate formate-lyase in Thermoanaerobacterium saccharolyticum JW/SL-YS485, Biotechnol. Biofuels. 8 (2015) 138. https://doi.org/10.1186/s13068-015-0304-1.

[27] C. Raynaud, P. Sarçabal, I. Meynial-Salles, C. Croux, P. Soucaille, Molecular characterization of the 1,3-propanediol (1,3-PD) operon of Clostridium butyricum, Proc. Natl. Acad. Sci. U. S. A. (2003). https://doi.org/10.1073/pnas.0734105100.

[28] D.G. Olson, M. Maloney, A.A. Lanahan, S. Hon, L.J. Hauser, L.R. Lynd, Identifying promoters for gene expression in Clostridium thermocellum, Metab. Eng. Commun. 2 (2015) 23–29. https://doi.org/10.1016/j.meteno.2015.03.002.

[29] A.M. Guss, D.G. Olson, N.C. Caiazza, L.R. Lynd, Dcm methylation is detrimental to plasmid transformation in Clostridium thermocellum, Biotechnol. Biofuels. 5 (2012) 30. https://doi.org/10.1186/1754-6834-5-30.

[30] J. Stern, S. Moraïs, Y. Ben-David, R. Salama, M. Shamshoum, R. Lamed, Y. Shoham, E.A. Bayer, I. Mizrahi, Assembly of synthetic functional cellulosomal structures onto the cell surface of Lactobacillus plantarum, a potent member of the gut microbiome, Appl. Environ.

Microbiol. 84 (2018) pii: e00282-18. https://doi.org/10.1128/AEM.00282-18.

[31] C. Gandini, L. Tarraran, D. Kalemasi, E. Pessione, R. Mazzoli, Recombinant Lactococcus lactis for efficient conversion of cellodextrins into L-lactic acid, Biotechnol. Bioeng. 114 (2017) 2807–2817. https://doi.org/10.1002/bit.26400.

[32] D.A. Ravcheev, X. Li, H. Latif, K. Zengler, S.A. Leyn, Y.D. Korostelev, A.E. Kazakov, P.S. Novichkov, A.L. Osterman, D.A. Rodionov, Transcriptional regulation of central carbon and energy metabolism in bacteria by redox-responsive repressor rex, J. Bacteriol. 194 (2012) 1145–1157. https://doi.org/10.1128/JB.06412-11.

[33] A.M. Williams-Rhaesa, N.K. Awuku, G.L. Lipscomb, F.L. Poole, G.M. Rubinstein, J.M. Conway, R.M. Kelly, M.W.W. Adams, Native xylose-inducible promoter expands the genetic tools for the biomass-degrading, extremely thermophilic bacterium Caldicellulosiruptor bescii, Extremophiles. 22 (2018) 629–638. https://doi.org/10.1007/s00792-018-1023-x.

[34] K. Willquist, E.W.J. van Niel, Lactate formation in Caldicellulosiruptor saccharolyticus is regulated by the energy carriers pyrophosphate and ATP, Metab. Eng. 12 (2010) 282–290. https://doi.org/10.1016/j.ymben.2010.01.001.

[35] F.O. Bryant, Characterization of the fructose 1,6-bisphosphate-activated, l(+lactate dehydrogenase from thermoanaerobacter ethanolicus, J. Enzyme Inhib. Med. Chem. 5 (1991) 235–248. https://doi.org/10.3109/14756369109080062.

[36] M. Özkan, E.I. Yilmaz, L.R. Lynd, G. Özcengiz, Cloning and expression of the Clostridium thermocellum L-lactate dehydrogenase gene in Escherichia coli and enzyme characterization, Can. J. Microbiol. 50 (2004) 845–851. https://doi.org/10.1139/w04-071.

[37] S. Hon, E.K. Holwerda, R.S. Worthen, M.I. Maloney, L. Tian, J. Cui, P.P. Lin, L.R. Lynd, D.G. Olson, Expressing the Thermoanaerobacterium saccharolyticum pforA in engineered

485 Clostridium thermocellum improves ethanol production, Biotechnol. Biofuels. 11 (2018)
486 242. https://doi.org/10.1186/s13068-018-1245-2.

487 [38] T. Rydzak, P.D. McQueen, O. V. Krokhin, V. Spicer, P. Ezzati, R.C. Dwivedi, D.
488 Shamshurin, D.B. Levin, J.A. Wilkins, R. Sparling, Proteomic analysis of Clostridium
489 thermocellum core metabolism: Relative protein expression profiles and growth phase-
490 dependent changes in protein expression, BMC Microbiol. 12 (2012) 214.
491 https://doi.org/10.1186/1471-2180-12-214.

492 [39] W. Xiong, P.P. Lin, L. Magnusson, L. Warner, J.C. Liao, P.C. Maness, K.J. Chou, CO2-
493 fixing one-carbon metabolism in a cellulose-degrading bacterium Clostridium thermocellum,
494 Proc. Natl. Acad. Sci. U. S. A. 113 (2016) 13180–13185.
495 https://doi.org/10.1073/pnas.1605482113.

496 [40] L. Copeland, A. Zammit, Kinetic properties of nad-dependent glyceraldehyde-3-phosphate
497 dehydrogenase from the host fraction of soybean root nodules, Arch. Biochem. Biophys. 312
498 (1994) 107–113. https://doi.org/10.1006/abbi.1994.1287.

499 [41] S. Even, C. Garrigues, P. Loubiere, N.D. Lindley, M. Cocaign-Bousquet, Pyruvate
500 Metabolism in Lactococcus lactis Is Dependent upon Glyceraldehyde-3-phosphate
501 Dehydrogenase Activity, Metab. Eng. 1 (1999) 198–205.
502 https://doi.org/10.1006/mben.1999.0120.

503 [42] L. Tian, S.J. Perot, D. Stevenson, T. Jacobson, A.A. Lanahan, D. Amador-Noguez, D.G.
504 Olson, L.R. Lynd, Metabolome analysis reveals a role for glyceraldehyde 3-phosphate
505 dehydrogenase in the inhibition of C. thermocellum by ethanol, Biotechnol. Biofuels. 10
506 (2017) 1–11. https://doi.org/10.1186/s13068-017-0961-3.

507 [43] J. Cui, D.G. Olson, L.R. Lynd, Characterization of the Clostridium thermocellum AdhE,
508 NfnAB, ferredoxin and Pfor proteins for their ability to support high titer ethanol production

509 in Thermoanaerobacterium saccharolyticum, Metab. Eng. 51 (2019) 32–42.
510 https://doi.org/10.1016/j.ymben.2018.09.006.

511 [44] R.A. Thompson, C.T. Trinh, Overflow metabolism and growth cessation in Clostridium
512 thermocellum DSM1313 during high cellulose loading fermentations, Biotechnol. Bioeng.
513 114 (2017) 2592–2604. https://doi.org/10.1002/bit.26374.

514 [45] H. Taguchi, H. Matsuzawa, T. Ohta, L-Lactate dehydrogenase from Thermus caldophilus
515 GK24, an extremely thermophilic bacterium., Eur. J. Biochem. 145 (1984) 283–290.

516

Table 1. *C. thermocellum* strains used in this study.

Strain	Description	SRA accession[1]	Reference
LL345	DSM 1313 Δ*hpt*	SRX872655	[22]
LL1111	LL345 Δ*adhE ldh*(S161R)	SRX744221	[20]
LL1147	LL345 Δ*hydG* Δ*ech*	SRX2141488	[23]
LL1624	LL345 P_{Thl}-*ldh*	SRX5676996	This study
LL1625	LL1147 P_{Thl}-*ldh*	SRX5678334	This study
LL1626	LL1147 P_{2638}-*ldh*	SRX5678333	This study
LL1640	LL345 P_{2638}-*ldh*	SRX6875981	This study

[1] Sequence Read Archive https://www.ncbi.nlm.nih.gov/sra

524 Table 2. Growth and fermentation profiles of the *C. thermocellum* strains. The data are the mean of three biological replicates ± standard deviation.

525 Light red and light green indicate values which are significantly ($p < 0.05$) lower and higher than data observed in the parent strain, respectively.

526 n.d., not detected.

527

Strain	Name	LL345	LL1624	LL1640	LL1111	LL1147	LL1625	LL1626
	Parent strain	LL345			LL1111	LL1147		
	ldh promoter	WT	P_{Thl}	P_{2638}	WT	WT	P_{Thl}	P_{2638}
Fermentation profile	growth rate (h^{-1})	0.5 ± 0.0	0.3 ± 0.0	0.4 ± 0.0	0.3 ± 0.1	0.4 ± 0.0	0.4 ± 0.0	0.3 ± 0.0
	final OD_{600nm}	1.6 ± 0.0	1.4 ± 0.0	1.4 ± 0.0	1.3 ± 0.0	1.3 ± 0.0	1.4± 0.0	1.2 ± 0.0
	Initial cellobiose (mM)	27.6 ± 0.4	27.0 ± 0.3	27.4 ± 0.5	27.0 ± 0.4	26.9 ± 0.3	26.7 ± 0.3	27.1 ± 0.1
	Consumed cellobiose (mM)	21.1 ± 0.4	21.4 ± 0.1	21.6 ± 0.5	17.8 ± 0.5	18.9 ± 0.3	19.6 ± 0.4	19.8 ± 0.2
	residual cellobiose (mM)	6.5 ± 0.3	5.6 ± 0.2	5.8 ± 0.2	9.2 ± 0.8	8.0 ± 0.1	7.1 ± 0.4	7.2 ± 0.3
	glucose (mM)	15.8 ± 1.1	14.6 ± 0.1	17.6 ± 0.6	18.3 ± 1.5	14.3 ± 0.6	10.3 ± 0.7	9.0 ± 0.5
	Residual hexose equivalent %	52.1 ± 1.2	47.7 ± 0.2	53.3 ± 2.1	67.8 ± 5.0	56.5 ± 1.4	45.9 ± 2.1	43.3 ± 1.6
	acetate (mM)	15.4 ± 1.5	9.1 ± 0.4	16.9 ± 2.4	8.6 ± 0.9	4.3 ± 0.4	4.1 ± 0.2	4.3 ± 0.0

	ethanol (mM)	7.6 ± 2.1	14.7 ± 0.2	5.8 ± 2.9	n.d.	8.5 ± 1.7	27.0 ± 1.8	29.7 ± 0.8
	formate (mM)	5.2 ± 0.2	4.6 ± 0.9	1.3 ± 0.2	0.1 ± 0.0	15.9 ± 0.0	14.2 ± 0.3	12.0 ± 0.3
	lactate (mM)	4.8 ± 0.2	11.1 ± 1.6	10.8 ± 0.4	21.5 ± 1.1	1.9 ± 0.2	0.6 ± 0.0	5.4 ± 0.6
	malate (mM)	1.6 ± 0.2	2.3 ± 0.2	3.1 ± 0.6	2.3 ± 0.3	1.0 ± 0.1	0.9 ± 0.0	0.5 ± 0.0
	pyruvate (mM)	0.1 ± 0.2	n.d.	n.d.	n.d.	1.8 ± 0.1	0.1 ± 0.1	0.0 ± 0.0
	succinate (mM)	-0.2 ± 0.0	-0.3 ± 0.0	-0.3 ± 0.0	-0.3 ± 0.1	0.0 ± 0.0	0.0 ± 0.1	-0.2 ± 0.0
	$Y_{acetate}$ (mol/mol cellobiose consumed)	1.2 ± 0.1	0.6 ± 0.0	1.3 ± 0.3	1.0 ± 0.2	0.4 ± 0.0	0.3 ± 0.0	0.3 ± 0.0
	$Y_{ethanol}$ (mol/mol cellobiose consumed)	0.6 ± 0.2	1.0 ± 0.0	0.4 ± 0.2	n.d.	0.7 ± 0.1	1.9 ± 0.1	1.9 ± 0.0
	$Y_{formate}$ (mol/mol cellobiose consumed)	0.4 ± 0.0	0.3 ± 0.1	0.1 ± 0.0	0.0 ± 0.0	1.4 ± 0.1	1.0 ± 0.0	0.8 ± 0.0
	$Y_{lactate}$ (mol/mol cellobiose consumed)	0.4 ± 0.0	0.8 ± 0.1	0.8 ± 0.1	2.5 ± 0.5	0.2 ± 0.0	0.0 ± 0.0	0.4 ± 0.0
	Y_{malate} (mol/mol cellobiose consumed)	0.1 ± 0.0	0.2 ± 0.0	0.2 ± 0.1	0.3 ± 0.1	0.1 ± 0.0	0.1 ± 0.0	0.0 ± 0.0
	$Y_{pyruvate}$ (mol/mol cellobiose consumed)	0.0 ± 0.0	0	n.d.	n.d.	0.2 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
	$Y_{succinate}$ (mol/mol cellobiose consumed)	-0.0 ± 0.0	-0.0 ± 0.0	-0.0 ± 0.0	-0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0

528

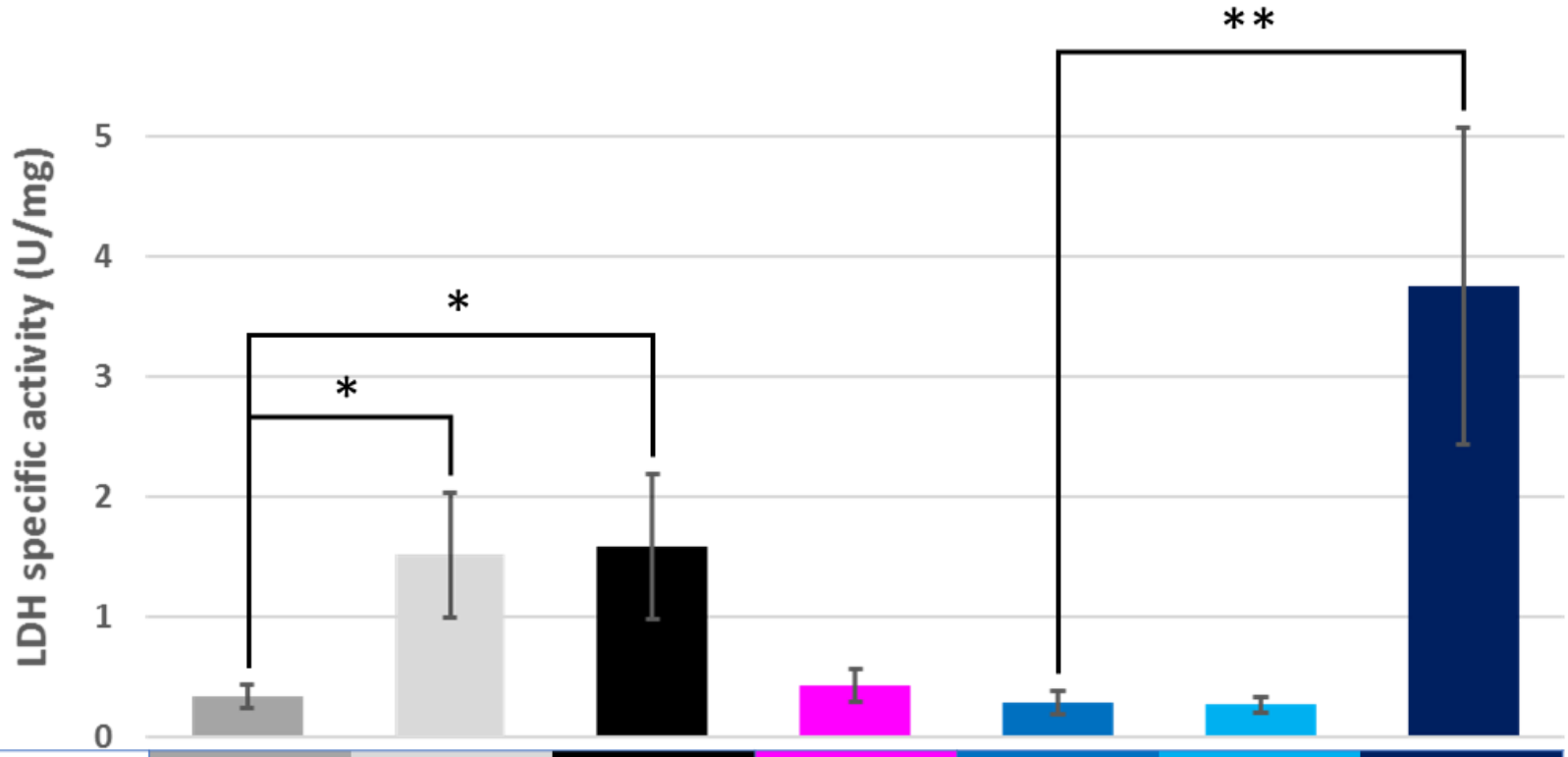
529

Figure legends

Figure 1. Lactate dehydrogenase (LDH) specific activity measured in the acellular crude extracts of *C. thermocellum* strains and LA yields (Y_{LA}) determined in the same strains. The data are the mean of three biological replicates ± standard deviation (for Y_{LA}, standard deviation is in parentheses). The * and ** labels indicate values that are significantly different ($p < 0.05$) from those observed in the corresponding parent strain. *, $p = 0.007$; ** $p = 0.0003$

Figure 2. Mutations found through genome sequencing of the *C. thermocellum* strains engineered in this study

Figure 1



Strain	LL345	LL1624	LL1640	LL1111	LL1147	LL1625	LL1626
Parent strain	LL345			LL1111	LL1147		
ldh promoter	WT	P_{Thl}	P_{2638}	WT	WT	P_{Thl}	P_{2638}
Y_{LA} (mol/mol cellobiose)	0.4 (0.0)	0.8 (0.1)	0.8 (0.1)	2.5 (0.5)	0.2 (0.0)	0.0 (0.0)	0.4 (0.0)

Figure 2

					Read fraction[2]			
Start Region[1]	Description	Type	Annotation name	Annotation description	LL1624	LL1640	LL1625	LL1626
4279	A --> -	Deletion	130 bp upstream of Clo1313_0005	cce:Ccel_0005 hypothetical protein	0,0	1,0	0,0	0,0
200106	A --> -, Ile127fs	Deletion	Clo1313_0185	phosphoribosyltransferase	0,0	1,0	0,0	0,0
279492	G --> -	Deletion	no CDS match		0,0	0,0	1,0	1,0
477437	C --> T, Leu38Phe	SNV	Clo1313_0422	ATP phosphoribosyltransferase catalytic region	1,0	0,0	0,0	0,0
494751	A --> -, Ile79fs	Deletion	Clo1313_0439	Sulfate transporter/antisigma-factor antagonist STAS	0,0	1,0	0,0	0,0
523339	A --> -, Lys28fs	Deletion	Clo1313_0478	type III restriction protein res subunit	0,0	0,0	0,0	0,0
618489	G --> A	SNV	288 bp upstream of Clo1313_0559	act:ACLA_088240 heat shock Hsp30-like protein, putative	1,0	0,0	0,0	0,0
677241	T --> C	SNV	no CDS match		0,4	0,0	0,0	0,0
677241	T --> T	SNV	no CDS match		0,6	0,0	0,0	0,0
724743	A --> G, Ile71Thr	SNV	Clo1313_0638	histone family protein DNA-binding protein	1,0	0,0	0,0	0,0
754780	ATTTAGTA --> -	Deletion	93 bp upstream of Clo1313_0663	dae:Dtox_2165 hypothetical protein	0,0	0,0	0,0	1,0
1058282	A --> -, Lys43fs	Deletion	Clo1313_0908	phosphoribosyltransferase	1,0	0,0	0,0	0,0
1193090	G --> T, Trp221Cys	SNV	Clo1313_1020	leucyl-tRNA synthetase	1,0	0,0	0,0	0,0
1198434	C --> T	SNV	Clo1313_1021	PKD domain containing protein	0,4	0,0	0,0	0,0
1198434	C --> C	SNV	Clo1313_1021	PKD domain containing protein	0,6	0,0	0,0	0,0
1232526	G --> T, Glu114*	SNV	Clo1313_1035	tRNA/rRNA methyltransferase (SpoU)	1,0	0,0	0,0	0,0
1271372	IS120	Insertion	no CDS match		0,0	1,0	0,0	0,0
1341055	A --> G	SNV	97 bp upstream of Clo1313_1122	metal-dependent phosphohydrolase HD sub domain	0,0	0,0	1,0	0,0
1370265	G --> A	SNV	45 bp upstream of Clo1313_1152	VanW family protein	1,0	0,0	0,0	0,0
1381635	P_{2638}	Insertion	ldh	Lactate/malate dehydrogenase	0,0	1,0	0,0	1,0
1381635	P_{Thl}	Insertion	ldh	Lactate/malate dehydrogenase	1,0	0,0	1,0	0,0
1438373	ISCth10	Insertion	89 bp upstream of Clo1313_1211	aspartate/glutamate/uridylate kinase	0,0	0,0	0,0	0,0
1570061	C --> A, Ala230Asp	SNV	Clo1313_1324	sigma-70 region 3 domain protein	0,0	0,0	1,0	0,0
1719993	T --> C, Val125Ala	SNV	Clo1313_1467	aminotransferase class I and II	0,0	0,0	1,0	0,0

1932207	IS120	Insertion	no CDS match		0,0	1,0	0,0	0,0
2314895	C --> T, Met85Ile	SNV	Clo1313_1970	Rhomboid family protein	1,0	0,0	0,0	0,0
2343123	- --> TATA	Insertion	14 bp upstream of Clo1313_1989	VTC domain	0,0	0,0	0,0	0,9
2624673	A --> T	SNV	no CDS match		1,0	0,0	0,0	0,0
2941043	C --> C	SNV	no CDS match		0,6	0,0	0,0	0,0
2941043	C --> T	SNV	no CDS match		0,4	0,0	0,0	0,0
2964040	G --> G	SNV	no CDS match		0,5	0,6	0,0	0,0
2988982	G --> A, Ala104Val	SNV	Clo1313_2549	peptidase S41	1,0	0,0	0,0	0,0
3001017	ISCth10	Insertion	Clo1313_2560	Carbamoyl-phosphate synthase L chain ATP-binding	0,0	0,0	1,0	0,0
3132643	C --> T	SNV	Clo1313_2666	rca:Rcas_1719 hypothetical protein	1,0	1,0	0,0	0,0
3151076	A --> C	SNV	132 bp upstream of Clo1313_2686	transposase IS200-family protein	0,0	0,0	0,0	0,5
3151076	A --> A	SNV	132 bp upstream of Clo1313_2686	transposase IS200-family protein	0,0	0,0	0,0	0,5
3301521	G --> T, Phe253Leu	SNV	Clo1313_2809	cpy:Cphy_2889 hypothetical protein	0,0	0,0	1,0	0,0
3359106	C --> T, Gly358Glu	SNV	Clo1313_2858	Carbohydrate binding family 6	0,0	1,0	0,0	0,0

[1] Start region is based on the coordinates from the *C. thermocellum* genome, Genbank accession number NC_017304.1

[2] Read fraction indicates the fraction of reads which support the presence of a given mutation. Read fractions > 0.95 indicate a mutation was called with high confidence. Lower read fraction values can result from a variety of causes, including duplicated genome regions, low read counts, and sequencing instrument noise.

Supplementary online material for

Construction of lactic acid overproducing *Clostridium thermocellum* through enhancement of lactate dehydrogenase expression

Mazzoli R.[a,b], Olson D.G.[b], Lynd L.R.[b]

[a]Structural and Functional Biochemistry. Laboratory of Proteomics and Metabolic Engineering of Prokaryotes. Department of Life Sciences and Systems Biology. University of Torino. Via Accademia Albertina 13. 10123 Torino. Italy.

[b]Thayer School of Engineering, Dartmouth College, 14 Engineering Drive, Hanover, NH 03755, USA.

Corresponding Author: Roberto Mazzoli, roberto.mazzoli@unito.it

Supplementary Table S1. Sequence of the gBlocks used to construct plasmids for functional replacement of the *ldh* promoter with the promoter of the thiolase (*thlA*) (P_{Thl}, gBlock XD904) of *Clostridium acetobutylicum* or the promoter of Clo1313_2638 (P_{2638}, gBlock XD902) of *C. thermocellum* DSM1313. Underlined sequences correspond to the sequence of the promoters.

gBlock name (lenght)	Purpose	Sequence (5'-3')
XD901 (540 bp)	5' flank	TAGGCGTATCACGAGGCGATCTTTTTCCCCAAAACTTCCGCAACAGTCTCCTTTGTAAGGTCATCCTGCGTGGGGCCGAGTCCTCCGGTCATAATAAC AAGGTCGCACCTTTCCAAAGCTGCAAGAAGACATTTTTTCAGCCGAACGGAATTGTCCCCCACCACACTGTGATAATACACATTCACACCAATGTCAT TGAGCCTTTTGGATATATACTGGGCATTGGTATTTGCTATCTGCCCCATTAAAAGCTCGGTTCCAACCGCTAATATCTCCGCATTCATATTGAAAGACC CCTTAAATTTAAACTTTTTGTAACTTATTATATCAATTAGTGTTATAAAATAAAAGGGAAAAAGAATTAAAATCAAAGGTTTCAAGAGCAGCCGTATC ACCCGTAAAAGTTTCAGCCGATTCAACCTTTTTACACATAAAACTTTCAAAAATTGATGACTTACAATTATCAAGTAGGATATAATATTACTAATGCTA AACAGTTATTGATAAAGGAGGAAGGAATATCGTGGGAATAGGCATGGA
XD902 (1249 bp)	insertion of promoter P_{2638}	TACCTGGCCCAGTAGTTCAGCTTTTTCCCCAAAACTTCCGCAACAGTCTCCTTTGTAAGGTCATCCTGCGTGGGGCCGAGTCCTCCGGTCATAATAAC AAGGTCGCACCTTTCCAAAGCTGCAAGAAGACATTTTTTCAGCCGAACGGAATTGTCCCCCACCACACTGTGATAATACACATTCACACCAATGTCAT TGAGCCTTTTGGATATATACTGGGCATTGGTATTTGCTATCTGCCCCATTAAAAGCTCGGTTCCAACCGCTAATATCTCCGCATTCATATTGAAAGACC CCTTAAATTTAAACTTTTTGTAACTTATTATATCAATTAGTGTTATAAAATAAAAGGGAAAAAGAATTAAAATCAAAGGTTTCAAGAGCAGCCGTATC ACCCGTAAAAGTTTCAGCCGATTCAACCTTTTTACACATAAAACTTTCAAAAATTGATGACTTACAATTATCAAGTAGGATATAATATTACTAATGCTA AACAGTTATTGATAAAGGAGGAAGGAATGATAAACAAAGGACGGTTCAGGGCTTCTGCTCATCCTACTCTGCATTGTAAAAAGGTAGGATGAATTTT TATTTTTAATCTTATTGAAAAAAATTTTTGAAAATCGGTTTTATTAAAAAAAAGTGGGTATATTTATAATAGTCAATTGATTGGTTAAAAAAATTTAAA TAAGCAAACAGAATAATAACAAAAGTAAGGAGGAATTTGTTATGAACAATAACAAAGTAATTAAAAAAGTAACCGTAGTTGGTGCAGGCTTTGTAG

		GTTCCACCACAGCTTATACATTGATGCTCAGCGGACTTATATCTGAAATTGTACTGATAGACATAAATGCAAAAAAAGCCGACGGAGAAGTCATGGA CTTAAATCACGGCATGCCTTTTGTAAGGCCCGTTGAAATTTATCGTGGTGACTACAAAGACTGTGCCGGATCCGACATAGTAATCATTACCGCCGGTG CCAACCAAAAAGAAGGCGAAACGAGAATAGATCTTGTTAAAAGAAACACGGAAGTATTCAAAAATATCATAAATGAAATTGTAAAGTACAACAACG ATTGTATTCTTCTGGTAGTCACAAATCCGGTGGATATTTTAACCTATGTAACTTACAAACTATCCGGATTCCCGAAAAACAAAGTAATAGGTTCCGGA ACGGTTTTGGACACAGCCAGGTTCCGTTATCTTTTAAGCGAACATGTAAAAGTGGACTGCTAATAGTAGTGAAAAA
XD904 (1203 bp)	insertion of promoter P_{Thl}	TACCTGGCCCAGTAGTTCAGCTTTTTCCCCAAAACTTCCGCAACAGTCTCCTTTGTAAGGTCATCCTGCGTGGGGCCGAGTCCTCCGGTCATAATAAC AAGGTCGCACCTTTCCAAAGCTGCAAGAAGACATTTTTTCAGCCGAACGGAATTGTCCCCCACCACACTGTGATAATACACATTCACACCAATGTCAT TGAGCCTTTTGGATATATACTGGGCATTGGTATTTGCTATCTGCCCCATTAAAAGCTCGGTTCCAACCGCTAATATCTCCGCATTCATATTGAAAGACC CCTTAAATTTAAACTTTTTGTAACTTATTATATCAATTAGTGTTATAAAATAAAAGGGAAAAAGAATTAAAATCAAAGGTTTCAAGAGCAGCCGTATC ACCCGTAAAAGTTTCAGCCGATTCAACCTTTTTACACATAAAACTTTCAAAAATTGATGACTTACAATTATCAAGTAGGATATAATATTACTAATGCTA AACAGTTATTGATAAAGGAGGAAGGAATTCGACTTTTTAACAAAATATATTGATAAAAATAATAATAGTGGGTATAATTAAGTTGTTAGAGAAAACG TATAAATTAGGGATAAACTATGGAACTTATGAAATAGATTGAAATGGTTTATCTGTTACCCCGTAGGATCCAGAATTTAAAAGGAGGGATTAAAATG AACAATAACAAAGTAATTAAAAAAGTAACCGTAGTTGGTGCAGGCTTTGTAGGTTCCACCACAGCTTATACATTGATGCTCAGCGGACTTATATCTGA AATTGTACTGATAGACATAAATGCAAAAAAAGCCGACGGAGAAGTCATGGACTTAAATCACGGCATGCCTTTTGTAAGGCCCGTTGAAATTTATCGT GGTGACTACAAAGACTGTGCCGGATCCGACATAGTAATCATTACCGCCGGTGCCAACCAAAAAGAAGGCGAAACGAGAATAGATCTTGTTAAAAGA AACACGGAAGTATTCAAAAATATCATAAATGAAATTGTAAAGTACAACAACGATTGTATTCTTCTGGTAGTCACAAATCCGGTGGATATTTTAACCTA TGTAACTTACAAACTATCCGGATTCCCGAAAAACAAAGTAATAGGTTCCGGAACGGTTTTGGACACAGCCAGGTTCCGTTATCTTTTAAGCGAACATG TAAAAGTGGACTGCTAATAGTAGTGAAAAA